

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Boodhoo, Amechand  
Seehra, Jasbir  
Shaw, Gray  
Sako, Dianne
- (ii) TITLE OF INVENTION: HIGHLY PURIFIED MOCARHAGIN, A COBRA VENOM  
PROTEASE, POLYNUCLEOTIDES ENCODING SAME AND  
RELATED PROTEASES, AND THERAPEUTIC USES  
THEREOF
- (iii) NUMBER OF SEQUENCES: 22
- (iv) CORRESPONDENCE ADDRESS:  
(A) ADDRESSEE: Genetics Institute, Inc.  
(B) STREET: 87 CambridgePark Drive  
(C) CITY: Cambridge  
(D) STATE: Massachusetts  
(E) COUNTRY: USA  
(F) ZIP: 02140
- (v) COMPUTER READABLE FORM:  
(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:  
(A) APPLICATION NUMBER:  
(B) FILING DATE:  
(C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:  
(A) NAME: Brown, Scott A.  
(B) REGISTRATION NUMBER: 32,724  
(C) REFERENCE/DOCKET NUMBER: GI5293A
- (ix) TELECOMMUNICATION INFORMATION:  
(A) TELEPHONE: (617) 498-8224  
(B) TELEFAX: (617) 876-5851

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Thr	Asn	Thr	Pro	Glu	Gln	Asp	Arg	Tyr	Leu	Gln	Ala	Lys	Lys	Tyr	Ile
1				5					10					15	
Glu	Phe	Tyr	Val	Val	Val	Asp	Asn	Val	Met	Tyr	Arg	Lys	Tyr		
			20					25					30		

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 48 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Thr	Asn	Thr	Pro	Glu	Gln	Asp	Arg	Tyr	Leu	Gln	Ala	Lys	Lys	Tyr	Ile
1				5					10					15	
Glu	Phe	Tyr	Val	Val	Val	Asp	Asn	Val	Met	Tyr	Arg	Lys	Tyr	Thr	Gly
			20					25					30		
Lys	Leu	His	Val	Ile	Thr	Xaa	Xaa	Val	Tyr	Glu	Met	Asn	Ala	Leu	Asn
			35					40					45		

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Glu	Ala	Thr	Glu	Tyr	Glu	Tyr	Leu	Asp	Tyr	Asp	Phe	Leu	Pro	Glu
1			5				10					15		

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Gln	Ala	Thr	Glu	Tyr	Glu	Tyr	Leu	Asp	Tyr	Asp	Phe	Leu	Pro	Glu
1			5				10					15		

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2050 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 78..1940

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

AGTCAATAGG AGAAGAGCTC AGGTTGGCTT GGAAGCAGAA AGAGATTCCT GTCCACCACT 60  
CCAATCCAGG CTCCAAAATG ATCCAAGCTC TCTTGGTAGC TATATGCTTA GCGGTTTTTC 120  
CATATCAAGG GAGCTCTATA ATCCTGGAAT CCGGGAATGT TAATGATTAT GAAGTAGTGT 180  
ATCCACAAAA AGTGCCTGCA TTGTCCAAAG GAGGAGTTCA GAATCCTCAG CCAGAGACCA 240  
AGTATGAAGA TACAATGCAA TATGAATTTC ACGTGAACGG AGAGCCAGTG GTCCTTCACT 300  
TAGAAAGAAA TAAAGGACTT TTTTCAGAAG ATTACACTGA AACTCATTAT GCCCCTGATG 360  
GCAGAGAAAT TACAACAAGC TCTCCAGTTC AGGATCACTG CTATTATCAT GGTACATTC 420  
AGAATGAAGC TGA CTCAAGT GCAGTCATCA GTGCATGTGA TGGCTTGAAA GGACATTTCA 480  
AGCATCAAGG GGAGACATAC TTTATTGAGC CCTTGAGAGT TTCTGACAGT GAAGCCCATG 540  
CAATATACAA AGATGAAAAT GTAGAAGAAG AGGAAGAGAT CCCCCAAATC TGTGGGGTTA 600  
CCCAGACTAC TTGGGAATCA GATGAGCCGA TTGAAAAGTC CTCTCAGTTA ACTAATACTC 660  
CTGAACAAGA CAGGTACTTG CAGGCCAAAA AATACATCGA GTTTTACGTG GTTGTGGACA 720  
ATGTAATGTA CMGRAAATAC ACCGGCAAGT TACATGTTAT AACAAGAAGA GTATATGAAA 780  
TGGTCAACGC TTAAATACG ATGTACAGAC GTTTGAATTT TCACATAGCA CTGATTGGCC 840  
TAGAAATTTG GTCCAACGGA AATGAGATTA ATGTGCAATC AGACGTGCAG GCCACTTTGG 900  
ACTTATTTGG AGAATGGAGA GAAAATAAAT TGCTGCCACG CAAAAGGAAT GATAATGCTC 960  
AGTTACTCAC GAGCACTGAG TTCAATGGAA CTACTACAGG ACTTG GTTAC ATAGGCTCCC 1020  
TCTGTAGTCC GAAGAAATCT GTGGCAGTTG TTCAGGATCA TAGCAAAAGC ACAAGCATGG 1080  
TGGCAATTAC AATGGCCCAT CAGATGGGTC ATAATCTGGG CATGAATGAT GACAGAGCTT 1140  
CCTGTACTTG TGGTTCTAAC AAATGCATTA TGTCTACAAA ATATTATGAA TCTCTTTCTG 1200

AGTTCAGCTC TTGTAGTGTC CAGGAACATC GGGAGTATCT TCTTAGAGAC AGACCACAAT 1260  
 GCATTCTCAA CAAACCCTCG CGCAAAGCTA TTGTTACACC TCCAGTTTGT GGAAATTACT 1320  
 TTGTGGAGCG GGGAGAAGAA TGTGACTGTG GCTCTCCTGA GGATTGTCAA AATACCTGCT 1380  
 GTGATGCTGC AACTTGTAAG CTGCAACATG AGGCACAGTG TGACTCTGGA GAGTGTGTG 1440  
 AGAAATGCAA ATTTAAGGGA GCAGGAGCAG AATGCCGGGC AGCAAAGAAT GACTGTGACT 1500  
 TTCCTGAACT CTGCACTGGC CGATCTGCTA AGTGTCCCAA GGACAGCTTC CAGAGGAATG 1560  
 GACATCCATG CCAAACAAC CAAGGTTACT GCTACAATGG GACATGTCCC ACCTTGACAA 1620  
 ACCAATGTGC TACTCTCTGG GGGCCAGGTG CAAAATGTC TCCAGGTTTA TGTTTTATGT 1680  
 TGAAGTGAA TGCCCGAAGT TGTGGCTTGT GCAGAAAGGA AAATGGCAGA AAGATTCTAT 1740  
 GTGCAGCAA GGATGTAAAG TGTGGCAGGT TATTTTGCAA AAAGAAAAAC TCGATGATAT 1800  
 GCCACTGCCC ACTCCATCAA AGGACCCAAA TTATGGAATG GTTGCACCTG GAACAAAATG 1860  
 TGGAGTAAA AAGGTGTGCA GAAACAGGCA ATGTGTAAA GTATAGACAG CCAACTGATC 1920  
 AAGCACTGCT TCTCTCAATT TGATTTTGA GATCCTCCTT CCAGAAGGCT TTCCTCAAGT 1980  
 CCAAAGAGAC CCATCTGTCT TTATCCTACT AGTAAATCAC TCTTAGCTTT CAAAAAAAAA 2040  
 AAAAGTCGAC 2050

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 621 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Ile Gln Ala Leu Leu Val Ala Ile Cys Leu Ala Val Phe Pro Tyr





500	505	510
Leu Thr Asn Gln Cys Ala Thr Leu Trp Gly Pro Gly Ala Lys Met Ser		
515	520	525
Pro Gly Leu Cys Phe Met Leu Asn Trp Asn Ala Arg Ser Cys Gly Leu		
530	535	540
Cys Arg Lys Glu Asn Gly Arg Lys Ile Leu Cys Ala Ala Lys Asp Val		
545	550	555 560
Lys Cys Gly Arg Leu Phe Cys Lys Lys Lys Asn Ser Met Ile Cys His		
565	570	575
Cys Pro Leu His Gln Arg Thr Gln Ile Met Glu Trp Leu His Leu Glu		
580	585	590
Gln Asn Val Glu Leu Lys Arg Cys Ala Glu Thr Gly Asn Val Leu Lys		
595	600	605
Tyr Arg Gln Pro Thr Asp Gln Ala Leu Leu Leu Ser Ile		
610	615	620

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2297 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GTCGACCA GT CAACAGGAGA AAAGCTCAGG TTGGCTTGGA AGCAGAAAGA GATTCCTGTC	60
CACCAGTCCA ATCCAGGCTC CAAAATGATC CAAGCTCTCT TGGTAATTAT ATGCTTAGCG	120
GTTTTTCAT ATCAAGGGAG CTCTATAATC CTGGAATCTG GGAATGTAA TGATTATGAA	180
GTTGTGTATC CACAAAAAGT GCCTGCATTG CTCAAAGGAG GAGTTCAGAA TCCTCAGCCA	240



GAGACCAAGT ATGAAGATAC AATGCAATAT GAATTTCAAG TGAATGGAGA GCCAGTAGTC	300
CTTCACTTAG AAAGAAATAA AGGACTTTTT TCAGAAGATT ACACTGAAAC TCATTATGCC	360
CCTGATGGCA GAGAAATTAC AACAGCCCT CCGTTCAGG ATCACTGCTA TTATCATGGT	420
TACATTCAGA ATGAAGCTGA CTCAAGTGCA ATCATCAGTG CATGTGATGG CTTGAAAGGA	480
CATTTCAAGC ATCAAGGGGA GACATACTTT ATTGAGCCCT TGAAGCTTTT CGACAGTGAA	540
TCTCATGCAA TCTACAAAGA TGAAAATGTA GAAAACGAGG ATGAGACCCC CGAAACCTGT	600
GGGGTAACCG AGACTACTTG GGAGTCAGAT GAGTCCATCG AAAAGACCTC TCAGTTAACT	660
AACACTCCTG AACAGACGG GTACTTGCAG GCCAAAAAAT ACATCGAGTT TTACGTGGTT	720
GTGGACAACA GAATGTACAG GTATTACAAA CGCAATGAAC CTGCTATAAA AAGAAGAGTA	780
TATGAAATGG TCAACGCTGT AAATACGAAG TACAGACCTT TGAAAATTCA CATAACACTG	840
ATTGGCCTAG AAATTTGGTC CAACCATGAT AAGTTTGAAG TGAAGCCAGT AGCGGGTGCC	900
ACTTTGAAAT CATTTGAGA TTGGAGAGAA ACAGTTTTGC TGCCACGCAA AAGGAATGAT	960
AACGCTCAGT TACTCACGGG CATTGACTTC AATGGAAGT TTGTGGGAAT TGCTTACACG	1020
GGCACCTCT GCACTCAGAA TTCTGTAGCA GTTGTTTCAGG ATTATAACCG AAAAATAAGC	1080
ATGGTGGCAT CTACAATGGC CCATGAGTTG GGTGATAATC TGGGCCTTCA TCATGACGGA	1140
GCTTCCTGTA TTTGCAGTCT TAGACCATGC ATTATGTCTA AGGGACGGAC TGCACCTGCC	1200
TTTCAGTTCA GCTCTTGTAG TGTCCGGGAG TATCGGGAGT ATCTTCTTAG AGAAAGACCA	1260
CAATGCATTC TCAACAAACC CTTGAGCACA GATACTGTTT CACCTGCAAT TTGTGGAAAT	1320
TACTTTGTGG AGGAGGGAGA AGAATGTGAC TGTGGCTCTC CTGCGGATTG TCAAAGTGCC	1380
TGCTGCGATG CTGCAACTTG TTAGTTTAAG GGAGAAGAAG CAGAATGCCG GGCAGCAAAG	1440
GATGACTGTG ACTTGCCTGA ACTCTGCACT GGCCGATCTG TGGAGTGTCC CACGGACAGC	1500
TTGCAGAAGA ATGGACATCC ATGTCAAAAC AACAAAGGTT ACTGCTACAA TGGGGCATGT	1560
CCCACCTTCA CAAACCAATG TATTGCTCTC ATGGGGACAG ATTTTACTGT GAGTCCAGAT	1620





Lys Arg Asn Asp Asn Ala Gln Leu Leu Thr Gly Ile Asp Phe Asn Gly  
 290 295 300  
 Thr Val Val Gly Ile Ala Tyr Thr Gly Thr Leu Cys Thr Gln Asn Ser  
 305 310 315 320  
 Val Ala Val Val Gln Asp Tyr Asn Arg Lys Ile Ser Met Val Ala Ser  
 325 330 335  
 Thr Met Ala His Glu Leu Gly His Asn Leu Gly Leu His His Asp Gly  
 340 345 350  
 Ala Ser Cys Ile Cys Ser Leu Arg Pro Cys Ile Met Ser Lys Gly Arg  
 355 360 365  
 Thr Ala Pro Ala Phe Gln Phe Ser Ser Cys Ser Val Arg Glu Tyr Arg  
 370 375 380  
 Glu Tyr Leu Leu Arg Glu Arg Pro Gln Cys Ile Leu Asn Lys Pro Leu  
 385 390 395 400  
 Ser Thr Asp Thr Val Ser Pro Ala Ile Cys Gly Asn Tyr Phe Val Glu  
 405 410 415  
 Glu Gly Glu Glu Cys Asp Cys Gly Ser Pro Ala Asp Cys Gln Ser Ala  
 420 425 430  
 Cys Cys Asp Ala Ala Thr Cys  
 435

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2335 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GTGACCTCA GGTGGCTTG GAAGCAGAAA GAGATTCCTA TCCACCACTC CAATCCAGGC	60
TCCAAAATGA TCCAAGCTCT CTTGGTAGCT ATATGCTTAG CGGTTTTTCC ATATCAAGGG	120
AGCTCTATAA TCCTGGAATC CGGGAATGTT AATGATTATG AAGTAGTGTA TCCACAAAAA	180
GTGCCTGCAT TGTCCAAAGG AGGAGTTCAG AATCCTCAGC CAGAGACCAA GTATGAAGAT	240
ACAATGCAAT ATGAATTTC AATGAATGGA GAGCCAGTAG TCCTTCACCT AGAAAGAAAT	300
AAAGGACTTT TTTCAGAAGA TTACACTGAA ACTCATTATG CCTCTGATGG CAGAGAAATT	360
ACAACAAGCC CACTCGTTCA GGATCACTGC TATTATCATG GTTACATTCA GAATGAAGCT	420
GACTCAAGTG CAGTCATCAG TGCATGCGAT GGCTTGAAAG GACATTTTCA GCTTCAAGGG	480
GAGACATACT TTATTGAACC CTTGAAGATT TCCGACAGTG AAGCCCATGC AATCTACAAA	540
GATGAAAATG TAGAAAACGA GGATGAGACC CCCGAAACCT GTGGGGTAAC CGAGACTACT	600
TGGGAGTCAG ATGAGTCCAT TGAAAAGACC TCTCAGTTAA CTAACACTCC TGAACAAGAC	660
AGGTACTTGC AGGCCAAAAA ATACCTCGAG TTTTACGTGG TTGTGGACAA CATAATGTAC	720
AGGCATTACA AACCGGATAA ACCTGTTATA AAAAGAAGAG TATATGAAAT GATCAACACT	780
ATGAATATGG TGTACAATCG TTTGAATTTT CACATAGCAC TGATTGGCCT AGAAATTTGG	840
TCCAACAGAA ATGAGATTAA TGTGCAATCA GACGTGCAGG CCACTTTGGA CTTATTTGGA	900
GAATGGAGAG AAAAAAATT GCTGCCACGC AAAAGGAATG ATAATGCTCA GTTACTCAG	960
GGTATTGACT TCAAAGGAAC TCCTGTAGGA CTTGCTTACA TAGGTTCCAT CTGCAATCCG	1020
AAGAGTTCTG TAGCAGTTGT TCAGGATTAT AGCAGTAGAA CAAGCATGGT GGCAATTACA	1080
ATGGCCCATG AGATGGGTCA TAATATGGGC ATTCATCATG ACGGACCTTC CTGTACTTGT	1140
GGTTCTAACA AATGCGTTAT GTCTACAAGA CGTACTGAAC CTGCCTATCA GTTCAGCTCT	1200

TGTAGTGTCC GGGAACATCA GGAGTATCTT CTTAGAGACA GACCACAATG CATTTCTCAAC	1260
AAACCCTTGA GCACAGATAT TGTTCACCT CCAATTTGTG GAAATAACTT TGTGGAGGTG	1320
GGAGAAGAAT GTGACTGTGG CTCTCCTGCG GATTGTCAAA GTGCCTGCTG CGACGCTACA	1380
ACTTGTA AAC TACAACCTCA TGCACAGTGT GACTCCGAAG GGTGTTGTGA GAAATGCAAA	1440
TTTAAGGGAG CAGGAGCAGA ATGCCGGGCA GCAAAGGATG ACTGTGACTT GCCTGAACTC	1500
TGCACTGGCC AATCTGCTGA GTGTCCCACA GACATCTTCC AGAGGAATGG ACTTCCATGC	1560
CAAAACAACG AAGGTTACTG CTACAATGGG AAATGCCCCA TCATGACAAA CCAATGTATT	1620
GCTCTCCGGG GACCAGGTGT AAAAGTATCT CGAGATAGCT GTTTTACATT GAACCAGAGA	1680
ACCAGTGGTT GTGGCTTGTG CAGAAATGGAA TATGGTAGAA AGATTCCATG TGCAGCAAAG	1740
GATGTAAAGT GTGGCAGGTT ATTTTGCAAA AAGGGAAACT CGATGATATG CAACTGCTCA	1800
GTTCACCAC GTGACCCAAG TTATGGAATG GTTGAACCTG GAACAAAATG TGGAGATGGA	1860
ATGGTGTGCA GCAACAGGCA GTGTGTTGAT GTGAAGACAG CCTACTGATC AAGCACTGGC	1920
TTCTCTCAAT TTGATTTTGG AGGTCCTCCT TCCAGAACGC TTCCCTCAAG TCCAAAGAGA	1980
CCCATCTGTC TTTATCCTAC TAGTAAATCA CTCTTAGCTT TCAGATGGTA TCTAAAATTT	2040
AAAAATATTC TTCTCCATAA TTTAAACTGG TAATCTTTTG CTAAAATCAG ACCTTTTCCC	2100
TGCCACAAAG CTCCATGGTC ATGTACAGCA CCAAAGGCTT ATTTGCTAAC AAGAAAAAAA	2160
ATGGCCATTT TACTGTTTGC CAATTGCAAT TCACATTTAA TGCAACAAGC TCTGCCCTTT	2220
GAGCTGGCGT ACTCAAAGGC AATGCTCCCT CTCCCAAAT TATACGCTGG CTTTCCAAGA	2280
TGTAGCTGCT TCCATCAATA AACTATTCTC ATTCTGAAAA AAAAAAAAAAG TCGAC	2335

(2) INFORMATION FOR SEQ ID NO:10:

[illegible]

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Gly Leu Lys Gly His Phe Glu Leu Gln Gly Glu Thr Tyr Phe Ile Glu  
130 135 140

Pro Leu Lys Ile Ser Asp Ser Glu Ala His Ala Ile Tyr Lys Asp Glu  
145 150 155 160

Asn Val Glu Asn Glu Asp Glu Thr Pro Glu Thr Cys Gly Val Thr Glu  
165 170 175

Thr Thr Trp Glu Ser Asp Glu Ser Ile Glu Lys Thr Ser Gln Leu Thr  
180 185 190

Asn Thr Pro Glu Gln Asp Arg Tyr Leu Gln Ala Lys Lys Tyr Leu Glu  
195 200 205

Phe Tyr Val Val Val Asp Asn Ile Met Tyr Arg His Tyr Lys Arg Asp  
210 215 220

Lys Pro Val Ile Lys Arg Arg Val Tyr Glu Met Ile Asn Thr Met Asn  
225 230 235 240

Met Val Tyr Asn Arg Leu Asn Phe His Ile Ala Leu Ile Gly Leu Glu  
245 250 255

Ile Trp Ser Asn Arg Asn Glu Ile Asn Val Gln Ser Asp Val Gln Ala  
260 265 270

Thr Leu Asp Leu Phe Gly Glu Trp Arg Glu Lys Lys Leu Leu Pro Arg  
275 280 285

Lys Arg Asn Asp Asn Ala Gln Leu Leu Thr Gly Ile Asp Phe Lys Gly  
290 295 300

Thr Pro Val Gly Leu Ala Tyr Ile Gly Ser Ile Cys Asn Pro Lys Ser  
305 310 315 320

Ser Val Ala Val Val Gln Asp Tyr Ser Ser Arg Thr Ser Met Val Ala  
325 330 335

Ile Thr Met Ala His Glu Met Gly His Asn Met Gly Ile His His Asp  
340 345 350

Gly Pro Ser Cys Thr Cys Gly Ser Asn Lys Cys Val Met Ser Thr Arg



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100

355	360	365
Arg Thr Glu Pro Ala Tyr Gln Phe Ser Ser Cys Ser Val Arg Glu His		
370	375	380
Gln Glu Tyr Leu Leu Arg Asp Arg Pro Gln Cys Ile Leu Asn Lys Pro		
385	390	395 400
Leu Ser Thr Asp Ile Val Ser Pro Pro Ile Cys Gly Asn Asn Phe Val		
405	410	415
Glu Val Gly Glu Glu Cys Asp Cys Gly Ser Pro Ala Asp Cys Gln Ser		
420	425	430
Ala Cys Cys Asp Ala Thr Thr Cys Lys Leu Gln Pro His Ala Gln Cys		
435	440	445
Asp Ser Glu Gly Cys Cys Glu Lys Cys Lys Phe Lys Gly Ala Gly Ala		
450	455	460
Glu Cys Arg Ala Ala Lys Asp Asp Cys Asp Leu Pro Glu Leu Cys Thr		
465	470	475 480
Gly Gln Ser Ala Glu Cys Pro Thr Asp Ile Phe Gln Arg Asn Gly Leu		
485	490	495
Pro Cys Gln Asn Asn Glu Gly Tyr Cys Tyr Asn Gly Lys Cys Pro Ile		
500	505	510
Met Thr Asn Gln Cys Ile Ala Leu Arg Gly Pro Gly Val Lys Val Ser		
515	520	525
Arg Asp Ser Cys Phe Thr Leu Asn Gln Arg Thr Ser Gly Cys Gly Leu		
530	535	540
Cys Arg Met Glu Tyr Gly Arg Lys Ile Pro Cys Ala Ala Lys Asp Val		
545	550	555 560
Lys Cys Gly Arg Leu Phe Cys Lys Lys Gly Asn Ser Met Ile Cys Asn		
565	570	575

Cys Ser Val Ser Pro Arg Asp Pro Ser Tyr Gly Met Val Glu Pro Gly  
580 585 590

Thr Lys Cys Gly Asp Gly Met Val Cys Ser Asn Arg Gln Cys Val Asp  
595 600 605

Val Lys Thr Ala Tyr  
610

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2288 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

AGTCAACAGG AGAAAAGCTC AGGTTGGCTT GGAAGCAGAA AGAGATTCCT GTCCACCAGT	60
CCAATCCAGG CTCCAAAATG ATCCAAGCTC TCTTGGTAAT TATATGCTTA GTGGTTTTTC	120
CATATCAAGG GAGCTCTATA ATCCTGGAAT CTGGGAATGT TAATGATTAT GAAGTTGTGT	180
ATCCACAAAA AGTGCCTGCA TTGCTCAAAG GAGGAGTTCA GAATCCTCAG CCAGAGACCA	240
AGTATGAAGA TACAATGCAA TATGAATTTT AAGTGAATGG AGAGCCAGTA GTCCTTCACT	300
TAGAAAGAAA TAAAGGACTT TTTTCAGAAG ATTACACTGA AACTCATTAT GCCCCTGATG	360
GCAGAGAAAT TACAACAAGC CCTCCGGTTC AGGATCACTG CTATTATCAT GGTACATTC	420
AGAATGAAGC TGAATCAAGT GCAATCATCA GTGCATGTGA TGGCTTGAAA GGACATTTCA	480



CCATGTGCAG CAAAGGATGT AAAGTGTGGC AGATTATACT GCACAGAGAG AGACACAATG 1740

TCATGCCGAT TCCCACTGGA CCCAGATGGT GTTAATGGCT GAACCTGGAA CAAAATGTGG 1800

AGATGGAATG GTGTGCAGCA ACGGTCAGTG TGTTAATGTG CAGACAGCCT ACTGATCAAG 1860

CACTGGCTTC TCTCAATTG ATTTTGGAGA TCCTCCTTCC AGAACGCTTC CCTCAAGTCC 1920

AAAGAGACCC ATCTGTCTTT ATCCTACTAG TAAATCACTC TTAGCTTTCA GATGGTATCT 1980

AAAATTTATA ATATTTCTTC TCCATAATTT AAAGTGGTAA TCTTTTGCTA AAATCAGACC 2040

TTTCCCTGC CACAAAGCTC CATGGTCATG TACAGCACCA AAGGCTTATT TGCGAATAAG 2100

AAAAAAAAAT GGCAATTTTA CAGTTTCCCA ATTGCAATGC ATATTGAATG CAACAAGCTC 2160

TGCCCTTTGA GCTGGCGTAT TCAAAGGCAA TGCTCCCTCT CCCAAAATTA TACGCTGGCT 2220

TTCCAAGATG TAGCTGCTTC CATCAATAAA CTATTCTCAT TCTGAAAAAA AAAAAAAAAA 2280

AAGTCGAC 2288

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 521 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Met Ile Gln Ala Leu Leu Val Ile Ile Cys Leu Val Val Phe Pro Tyr  
 1 5 10 15

Gln	Gly	Ser	Ser	Ile	Ile	Leu	Glu	Ser	Gly	Asn	Val	Asn	Asp	Tyr	Glu			
																20	25	30
Val	Val	Tyr	Pro	Gln	Lys	Val	Pro	Ala	Leu	Leu	Lys	Gly	Gly	Val	Gln			
																35	40	45
Asn	Pro	Gln	Pro	Glu	Thr	Lys	Tyr	Glu	Asp	Thr	Met	Gln	Tyr	Glu	Phe			
																50	55	60
Gln	Val	Asn	Gly	Glu	Pro	Val	Val	Leu	His	Leu	Glu	Arg	Asn	Lys	Gly			
																65	70	75
Leu	Phe	Ser	Glu	Asp	Tyr	Thr	Glu	Thr	His	Tyr	Ala	Pro	Asp	Gly	Arg			
																85	90	95
Glu	Ile	Thr	Thr	Ser	Pro	Pro	Val	Gln	Asp	His	Cys	Tyr	Tyr	His	Gly			
																100	105	110
Tyr	Ile	Gln	Asn	Glu	Ala	Asp	Ser	Ser	Ala	Ile	Ile	Ser	Ala	Cys	Asp			
																115	120	125
Gly	Leu	Lys	Gly	His	Phe	Lys	His	Gln	Gly	Glu	Thr	Tyr	Phe	Ile	Glu			
																130	135	140
Pro	Leu	Lys	Leu	Phe	Asp	Ser	Glu	Ser	His	Ala	Ile	Tyr	Lys	Asp	Glu			
																145	150	155
Asn	Val	Glu	Asn	Glu	Asp	Glu	Thr	Pro	Glu	Thr	Cys	Gly	Val	Thr	Glu			
																165	170	175
Thr	Thr	Trp	Glu	Ser	Asp	Glu	Ser	Ile	Glu	Lys	Thr	Ser	Gln	Leu	Thr			
																180	185	190
Asn	Thr	Pro	Glu	Gln	Asp	Gly	Tyr	Leu	Gln	Ala	Lys	Lys	Tyr	Ile	Glu			
																195	200	205
Phe	Tyr	Val	Val	Val	Asp	Asn	Arg	Met	Tyr	Arg	Tyr	Tyr	Lys	Arg	Asn			
																210	215	220
Glu	Pro	Ala	Ile	Lys	Arg	Arg	Val	Tyr	Glu	Met	Val	Asn	Ala	Val	Asn			

225	230	235	240
Thr Tyr Arg Pro Leu Lys Ile His Ile Thr Leu Ile Gly Leu Glu Ile			
245	250	255	
Trp Ser Asn Asp Asp Lys Phe Glu Val Lys Pro Val Ala Gly Ala Thr			
260	265	270	
Leu Lys Ser Phe Arg Asp Trp Arg Glu Thr Val Leu Leu Pro Arg Lys			
275	280	285	
Arg Asn Asp Asn Ala Gln Leu Leu Thr Gly Ile Asp Phe Asn Gly Thr			
290	295	300	
Val Val Gly Ile Ala Tyr Thr Gly Thr Leu Cys Thr Gln Asn Ser Val			
305	310	315	320
Ala Val Val Gln Asp Tyr Asn Arg Lys Ile Ser Met Val Ala Ser Thr			
325	330	335	
Met Ala His Glu Leu Gly His Asn Leu Gly Leu His His Asp Gly Ala			
340	345	350	
Ser Cys Ile Cys Ser Leu Arg Pro Cys Ile Met Ser Lys Gly Arg Thr			
355	360	365	
Ala Pro Ala Phe Gln Phe Ser Ser Cys Ser Val Arg Glu Tyr Arg Glu			
370	375	380	
Tyr Leu Leu Arg Glu Arg Pro Gln Cys Ile Leu Asn Lys Pro Leu Ser			
385	390	395	400
Thr Asp Thr Val Ser Pro Ala Ile Cys Gly Asn Tyr Phe Val Glu Glu			
405	410	415	
Gly Glu Glu Cys Asp Cys Gly Ser Pro Ala Asp Cys Gln Ser Ala Cys			
420	425	430	
Cys Asp Ala Ala Thr Cys Lys Phe Lys Gly Glu Glu Ala Glu Cys Arg			
435	440	445	

Ala Ala Lys Asp Asp Cys Asp Leu Pro Glu Leu Cys Thr Gly Arg Ser  
450 455 460

Val Glu Cys Pro Thr Asp Ser Leu Gln Arg Asn Gly His Pro Cys Gln  
465 470 475 480

Asn Asn Lys Gly Tyr Cys Tyr Asn Gly Ala Cys Pro Thr Phe Thr Asn  
485 490 495

Gln Cys Ile Ala Leu Met Gly Thr Asp Phe Thr Val Ser Pro Asp Gly  
500 505 510

Cys Phe Asp Leu Asn Val Arg Gly Asn  
515 520

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2309 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GTCGACGTCA ACAGGAGAAA AGCTCAGGTT GGCTTGAAG CAGAAAGAGA TTCCTGTCCA	60
CCAGTCCAAT CCAGGCTCCA AAATGATCCA AGCTCTCTTG GTAATTATAT GCTTAGTGGT	120
TTTTCCATAT CAAGGGAGCT CTATAATCCT GGAATCTGGG AATGTTAATG ATTATGAAGT	180
TGTGTATCCA CAAAAAGTGC CTGCATTGCT CAAAGGAGGA GTTCAGAATC CTCAGCCAGA	240
GACCAAGTAT GAAGATACAA TGCAATATGA ATTTCAAGTG AATGGAGAGC CAGTAGTCCT	300





GCAGAGGAAT	GGACATCCAT	GTCAAAACAA	CAAAGGTTAC	TGCTACAATG	GGGCATGTCC	1560
CACCTTCACA	AACCAATGTA	TTGCTCTCAT	GGGGACAGAT	TTTACTGTGA	GTCCAGATGG	1620
ATGTTTTGAC	TTGAACGTGA	GAGGGAATGA	TGTAAGCCAC	TGCAGAAAGG	AAAATGGTGC	1680
AAAGATTCCA	TGTGCAGCAA	AGGATGTAAA	GTGTGGCAGG	TTATACTGCA	CAGAGAGAGA	1740
CACAATGTCA	TGCCGATTCC	CACTGGACCC	AGATGGTGTA	ATGGCTGAAC	CTGGAACAAA	1800
ATGTGGAGAT	GGAATGGTGT	GCAGCAACGG	TCAGTGTGTT	AATGTGCAGA	CAGCCTACTG	1860
ATCAAGCACT	GGCTTCTCTC	AATTTGATTT	TGGAGATCCT	CCTTCCAGAA	CGCTTCCCTC	1920
AAGTCCAAAG	AGACCCATCT	GTCTTTATCC	TACTAGTAAA	TCACTCTTAG	CTTTCAGATG	1980
GTATCTAAAA	TTTATAATAT	TTCTTCTCCA	TAATTTAAAC	TGGAATCTT	TTGCTAAAAA	2040
CAGACCTTTT	CCCTGCCACA	AAGCTCCATG	GTCATGTACA	GCACCAAAGG	CTTATTTGCG	2100
AATAAGAAAA	AAAAATGGCA	ATTTTACAGT	TTCCAATTG	CAATGCATAT	TGAATGCAAC	2160
AAGCTCTGCC	CTTTGAGCTG	GCGTATTCAA	AGGCAATGCT	CCCTCTCCCA	AAATTATACG	2220
CTGGCTTTCC	AAGATGTAGC	TGCTTCCATC	AATAAACTAT	TCTCATTCTG	AAAAAAAAAA	2280
AAAAAAAAAA	AAAAAAAAAA	AAAGTCGAC				2309

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 592 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met Ile Gln Ala Leu Leu Val Ile Ile Cys Leu Val Val Phe Pro Tyr  
1 5 10 15

Gln Gly Ser Ser Ile Ile Leu Glu Ser Gly Asn Val Asn Asp Tyr Glu  
20 25 30

Val Val Tyr Pro Gln Lys Val Pro Ala Leu Leu Lys Gly Gly Val Gln  
35 40 45

Asn Pro Gln Pro Glu Thr Lys Tyr Glu Asp Thr Met Gln Tyr Glu Phe  
50 55 60

Gln Val Asn Gly Glu Pro Val Val Leu His Leu Glu Arg Asn Lys Gly  
65 70 75 80

Leu Phe Ser Glu Asp Tyr Thr Glu Thr His Tyr Ala Pro Asp Gly Arg  
85 90 95

Glu Ile Thr Thr Ser Pro Pro Val Gln Asp His Cys Tyr Tyr His Gly  
100 105 110

Tyr Ile Gln Asn Glu Ala Asp Ser Ser Ala Ile Ile Ser Ala Cys Asp  
115 120 125

Gly Leu Lys Gly His Phe Lys His Gln Gly Glu Thr Tyr Phe Ile Glu  
130 135 140

Pro Leu Lys Leu Phe Asp Ser Glu Ser His Ala Ile Tyr Lys Asp Glu  
145 150 155 160

Asn Val Glu Asn Glu Asp Glu Thr Pro Glu Thr Cys Gly Val Thr Glu  
165 170 175

Thr Thr Trp Glu Ser Asp Glu Ser Ile Glu Lys Thr Ser Gln Leu Thr  
180 185 190

Asn Thr Pro Glu Gln Asp Gly Tyr Leu Gln Ala Lys Lys Tyr Ile Glu  
195 200 205

Phe Tyr Val Val Val Asp Asn Arg Met Tyr Arg Tyr Tyr Lys Arg Asn  
 210 215 220

Glu Pro Ala Ile Lys Arg Arg Val Tyr Glu Met Val Asn Ala Val Asn  
 225 230 235 240

Thr Lys Tyr Arg Pro Leu Lys Ile His Ile Thr Leu Ile Gly Leu Glu  
 245 250 255

Ile Trp Ser Asn Asp Asp Lys Phe Glu Val Lys Pro Val Ala Gly Ala  
 260 265 270

Thr Leu Lys Ser Phe Arg Asp Trp Arg Glu Thr Val Leu Leu Pro Arg  
 275 280 285

Lys Arg Asn Asp Asn Ala Gln Leu Leu Thr Gly Ile Asp Phe Asn Gly  
 290 295 300

Thr Val Val Gly Ile Ala Tyr Thr Gly Thr Leu Cys Thr Gln Asn Ser  
 305 310 315 320

Val Ala Val Val Gln Asp Tyr Asn Arg Lys Ile Ser Met Val Ala Ser  
 325 330 335

Thr Met Ala His Glu Leu Gly His Asn Leu Gly Leu His His Asp Gly  
 340 345 350

Ala Ser Cys Ile Cys Ser Leu Arg Pro Cys Ile Met Ser Lys Gly Arg  
 355 360 365

Thr Ala Pro Ala Phe Gln Phe Ser Ser Cys Ser Val Arg Glu Tyr Arg  
 370 375 380

Glu Tyr Leu Leu Arg Glu Arg Pro Gln Cys Ile Leu Asn Lys Pro Leu  
 385 390 395 400

Ser Thr Asp Thr Val Ser Pro Ala Ile Cys Gly Asn Tyr Phe Val Glu  
 405 410 415

Glu Gly Glu Glu Cys Asp Cys Gly Ser Pro Ala Asp Cys Gln Ser Ala

[illegible]

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1820 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GTCGACGACA TTTCAAGCAT CAAGGGGAGA CATACTTTAT TGAGCCCTTG AAGCTTTTCG	60
ACAGTGAATC CCATGCAATC TACAAAGATG AAAATGTAGA AAACGAGGAT GAGACCCCCG	120
AAACCTGTGG GGTAACCGAG ACTACTTGGG AGTCAGATGA GTCCATTGAA AAGACCTCTC	180
AGTTAACTAA CACTCCTGAA CAAGACGGGT ACTTGCAGGC CAAAAAATAC ATCGAGTTTT	240
ACGTGGTTGT GGACAACAGA ATGTACAGGT ATTACAAACG CAATGAACCT GCTATAAAAA	300
GAAGAGTATA TGAAATGGTC AACGCTGTAA ATACGAAGTA CAGACCTTG AAAATTCACA	360
TAACACTGAT TGGCCTAGAA ATTTGGTCCA ACGATGATAA GTTTGAAGTG AAGCCAGTAG	420
CGGGTGCCAC TTTGAAATCA TTTCGAGATT GGAGAGAAAC AGTTTTGCTG CCACGCAAAA	480
GGAATGATAA CGCTCAGTTA CTCACGGGCA TTGACTTCAA TGGAAGTGT GTGGGAATTG	540
CTTACACGGG CACCCTCTGC ACTCAGAATT CTGTAGCAGT TGTTCAAGGAT TATAACCGAA	600
AAATAAGCAT GGTGGCATCT ACAATGGCCC ATGAGTTGGG TCATAATCTG GGCATTTCATC	660
ATGACGGAGC TTCCTGTATT TGCAGTCTTA AACCATGCAT TATGTCTAAG GGACGGACTG	720
CACCTGCCTT TCAGTTCAGC TCTTGTAGTG TCCGGGAGTA TCGGGAGTAT CTTCTTAGAA	780
AAAGACCACA ATGCATTCTC AACAAACCCT TGAGCACAGA TATTGTTTCA CCTGCAATTT	840
GTGGAAATTA CTTTGTGGAG GAGGGAGAAG AATGTGACTG TGGCTCTCCT GCGGATTGTC	900
AAAGTGCCTG CTGCAATGCT GCAACTTGTA AGTTTAAGGG AGAAGAAGCA GAATGCCGGG	960
CAGCAAAGGA TGA CTGTGAC TTGCCTGAAC TCTGCACTGG CCGATCTGTG GAGTGTCCCA	1020

CGGACAGCTT GCAGAGGAAT GGACATCCAT GTCAAAACAA CAAAGGTTAC TGCTACAATG	1080
GGGCATGTCC CACCTTCACA AACCAATGTA TTGCTCTCAT GGGGACAGAT TTTACTGTGA	1140
GTCCAGATGG ATGTTTTGAC TTGAACGTGA GAGGGAATGA TGTAAGCCAC TGCAGAAAGG	1200
AAAATGGTGC AAAGATTCCA TGTGCAGCAA AGGATGTAAA GTGTGGCAGG TTATACTGCA	1260
CAGAGAGAAA CACAATGTCA TGCCGATTCC CACTGGACCC AGATGGTGTA ATGGCTGAAC	1320
CTGGAACAAA ATGTGGAGAT GGAATGGTGT GCAGCAACGG TCAGTGTGTT AATGTGCAGA	1380
CAGCCTACTG ATCAAGCACT GGCTTCTCTC AATTGATTT TGGAGATCCT CCTTCAGAA	1440
CGCTTCCCTC AAGTCCAAAG AGACCCATCT GTCTTTATCC TACTAGTAAA TCACTCTTAG	1500
CTTTCAGATG GTATCTAAAA TTTATAATAT TTCTTCTCCA TAATTTAAAC TGGTAATCTT	1560
TTGCTAAAAT CAGACCTTTT CCCTGCCACA AAGCTCCATG GTCATGTACA GTACCAAAGG	1620
CTTATTTGCT AACACGAAAA AAAATGGCCA TTTTACCGTT TGCCAATTGC AATTCACATT	1680
TAATGCAACA AGCTCTGCCC TTTGAGCTGG CGTATTCAAA GGCAATGCTC CCTCTCCCAA	1740
AATTATATGC TGGCTTTCCA AGATGTAGCT GCTTCCATCA ATAACTATT CTCATTCTGA	1800
AAAAAAAAAA AAAAGTCGAC	1820

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 462 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Arg	Arg	His	Phe	Lys	His	Gln	Gly	Glu	Thr	Tyr	Phe	Ile	Glu	Pro	Leu
1				5					10					15	
Lys	Leu	Phe	Asp	Ser	Glu	Ser	His	Ala	Ile	Tyr	Lys	Asp	Glu	Asn	Val
				20				25					30		
Glu	Asn	Glu	Asp	Glu	Thr	Pro	Glu	Thr	Cys	Gly	Val	Thr	Glu	Thr	Thr
	35						40					45			
Trp	Glu	Ser	Asp	Glu	Ser	Ile	Glu	Lys	Thr	Ser	Gln	Leu	Thr	Asn	Thr
	50					55						60			
Pro	Glu	Gln	Asp	Gly	Tyr	Leu	Gln	Ala	Lys	Lys	Tyr	Ile	Glu	Phe	Tyr
65					70					75				80	
Val	Val	Val	Asp	Asn	Arg	Met	Tyr	Arg	Tyr	Tyr	Lys	Arg	Asn	Glu	Pro
				85					90				95		
Ala	Ile	Lys	Arg	Arg	Val	Tyr	Glu	Met	Val	Asn	Ala	Val	Asn	Thr	Lys
				100				105					110		
Tyr	Arg	Pro	Leu	Lys	Ile	His	Ile	Thr	Leu	Ile	Gly	Leu	Glu	Ile	Trp
	115						120					125			
Ser	Asn	Asp	Asp	Lys	Phe	Glu	Val	Lys	Pro	Val	Ala	Gly	Ala	Thr	Leu
	130					135						140			
Lys	Ser	Phe	Arg	Asp	Trp	Arg	Glu	Thr	Val	Leu	Leu	Pro	Arg	Lys	Arg
145					150					155				160	
Asn	Asp	Asn	Ala	Gln	Leu	Leu	Thr	Gly	Ile	Asp	Phe	Asn	Gly	Thr	Val
				165					170				175		
Val	Gly	Ile	Ala	Tyr	Thr	Gly	Thr	Leu	Cys	Thr	Gln	Asn	Ser	Val	Ala
		180						185				190			
Val	Val	Gln	Asp	Tyr	Asn	Arg	Lys	Ile	Ser	Met	Val	Ala	Ser	Thr	Met
	195						200					205			

Ala His Glu Leu Gly His Asn Leu Gly Ile His His Asp Gly Ala Ser  
 210 215 220

Cys Ile Cys Ser Leu Lys Pro Cys Ile Met Ser Lys Gly Arg Thr Ala  
 225 230 235 240

Pro Ala Phe Gln Phe Ser Ser Cys Ser Val Arg Glu Tyr Arg Glu Tyr  
 245 250 255

Leu Leu Arg Lys Arg Pro Gln Cys Ile Leu Asn Lys Pro Leu Ser Thr  
 260 265 270

Asp Ile Val Ser Pro Ala Ile Cys Gly Asn Tyr Phe Val Glu Glu Gly  
 275 280 285

Glu Glu Cys Asp Cys Gly Ser Pro Ala Asp Cys Gln Ser Ala Cys Cys  
 290 295 300

Asn Ala Ala Thr Cys Lys Phe Lys Gly Glu Glu Ala Glu Cys Arg Ala  
 305 310 315 320

Ala Lys Asp Asp Cys Asp Leu Pro Glu Leu Cys Thr Gly Arg Ser Val  
 325 330 335

Glu Cys Pro Thr Asp Ser Leu Gln Arg Asn Gly His Pro Cys Gln Asn  
 340 345 350

Asn Lys Gly Tyr Cys Tyr Asn Gly Ala Cys Pro Thr Phe Thr Asn Gln  
 355 360 365

Cys Ile Ala Leu Met Gly Thr Asp Phe Thr Val Ser Pro Asp Gly Cys  
 370 375 380

Phe Asp Leu Asn Val Arg Gly Asn Asp Val Ser His Cys Arg Lys Glu  
 385 390 395 400

Asn Gly Ala Lys Ile Pro Cys Ala Ala Lys Asp Val Lys Cys Gly Arg  
 405 410 415

Leu Tyr Cys Thr Glu Arg Asn Thr Met Ser Cys Arg Phe Pro Leu Asp



420

425

430

Pro Asp Gly Val Met Ala Glu Pro Gly Thr Lys Cys Gly Asp Gly Met

435

440

445

Val Cys Ser Asn Gly Gln Cys Val Asn Val Gln Thr Ala Tyr

450

455

460

## (2) INFORMATION FOR SEQ ID NO:17:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2359 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GTCGACCTCA GGTGGCTTG GAAGCAGAAA GAGATTCCTA TCCACCACTC CAATCCAGGC	60
TCCAAAATGA TCCAAGCTCT CTTGGTAGCT ATATGCTTAG CGGTTTTTCC ATATCAAGGG	120
AGCTCTATAA TCCTGGAATC CGGGAATGTT AATGATTATG AAGTAGTGTA TCCACAAAAA	180
GTGCCTGCAT TGTCCAAAGG AGGAGTTCAG AATCCTCAGC CAGAGACCAA GTATGAAGAT	240
ACAATGCAAT ATGAATTTCA AGTGAATGGA GAGCCAGTAG TCCTTCACCT AGAAAGAAAT	300
AAAGGACTTT TTTCAGAAGA TTACACTGAA ACTCATTATG CCTCTGATGG CAGAGAAATT	360
ACAACAAGCC CACTCGTTCA GGATCACTGC TATTATCATG GTTACATTCA GAATGAAGCT	420
GACTCAAGTG CAGTCATCAG TGCATGCGAT GGCTTGAAAG GACATTTCTGA GCTTCAAGGG	480
GAGACATACT TTATTGAACC CTTGAAGATT TCCGACAGTG AAGCCCATGC AATCTACAAA	540
GATGAAAATG TAGAAAACGA GGATGAGACC CCCGAAACCT GTGGGGTAAC CGAGACTACT	600
TGGGAGTCAG ATGAGTCCAT TGAAGAGACC TCTCAGTTAG ACGACGACGA CAAGCGGCCG	660

CCAAC TAACA CTCTGAACA AGACAGGTAC TTGCAGGCCA AAAAATACCT CGAGTTTAC	720
GTGGTTGTGG ACAACATAAT GTACAGGCAT TACAAACGCG ATAAACCTGT TATAAAAAGA	780
AGAGTATATG AAATGATCAA CACTATGAAT ATGGTGTACA ATCGTTTGAA TTTTCACATA	840
GCACTGATTG GCCTAGAAAT TTGGTCCAAC AGAAATGAGA TTAATGTGCA ATCAGACGTG	900
CAGGCCACTT TGGACTTATT TGGAGAATGG AGAGAAAAAA AATTGCTGCC ACGCAAAAGG	960
AATGATAATG CTCAGTTACT CACGGGTATT GACTTCAAAG GAACTCCTGT AGGACTTGCT	1020
TACATAGGTT CCATCTGCAA TCCGAAGAGT TCTGTAGCAG TTGTT CAGGA TTATAGCAGT	1080
AGAACAAGCA TGGTGGCAAT TACAATGGCC CATGAGATGG GTCATAATAT GGGCATT CAT	1140
CATGACGGAC CTTCCTGTAC TTGTGGTTCT AACAAATGCG TTATGTCTAC AAGACGTACT	1200
GAACCTGCCT ATCAGTTCAG CTCTTG TAGT GTCCGGGAAC ATCAGGAGTA TCTTCTTAGA	1260
GACAGACCAC AATGCATTCT CAACAAACCC TTGAGCACAG ATATTGTTTC ACCTCCAATT	1320
TGTGGAAATA ACTTTGTGGA GGTGGGAGAA GAATGTGACT GTGGCTCTCC TCGCGATTGT	1380
CAAAGTGCCT GCTGCGACGC TACAAC TTGT AACTTACAAC CTCATGCACA GTGTGACTCC	1440
GAAGGGTGTT GTGAGAAATG CAAATTTAAG GGAGCAGGAG CAGAA TGCCG GGCAGCAAAG	1500
GATGACTGTG ACTTGCTGA ACTCTGCACT GGCCAATCTG CTGAGTGTCC CACAGACATC	1560
TTCCAGAGGA ATGGACTTCC ATGCCAAAAC AACGAAGGTT ACTGCTACAA TGGGAAATGC	1620
CCCATCATGA CAAACCAATG TATTGCTCTC CGGGGACCAG GTGTAAAAGT ATCTCGAGAT	1680
AGCTGTTTTA CATTGAACCA GAGAACCAGT GGTGTGGCT TGTGCAGAAAT GGAATATGGT	1740
AGAAAGATTC CATGTGCAGC AAAGGATGTA AAGTGTGGCA GGTATT TTG CAAAAAGGGA	1800
AACTCGATGA TATGCAACTG CTCAGTTTCA CCACGTGACC CAAGTTATGG AATGGTTGAA	1860
CCTGGAACAA AATGTGGAGA TGGAATGGTG TGCAGCAACA GGCAGTGTGT TGATGTGAAG	1920
ACAGCCTACT GATCAAGCAC TGGCTTCTCT CAATTTGATT TTGGAGGTCC TCCTTCCAGA	1980
ACGCTTCCCT CAAGTCCAAA GAGACCCATC TGTCTTTATC CTACTAGTAA ATCACTCTTA	2040

GCTTTCAGAT GGTATCTAAA ATTTAAAATA TTTCTTCTCC ATAATTTAAA CTGGTAATCT 2100  
TTTGCTAAAA TCAGACCTTT TCCCTGCCAC AAAGCTCCAT GGTCATGTAC AGCACCAAAG 2160  
GCTTATTTGC TAACAAGAAA AAAAATGGCC ATTTTACTGT TTGCCAATTG CAATTCACAT 2220  
TTAATGCAAC AAGCTCTGCC CTTTGAGCTG GCGTACTCAA AGGCAATGCT CCCTCTCCCA 2280  
AAATTATACG CTGGCTTTCC AAGATGTAGC TGCTTCCATC AATAAACTAT TCTCATTCTG 2340  
AAAAAAAAAA AAAGTCGAC 2359

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 621 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Met	Ile	Gln	Ala	Leu	Leu	Val	Ala	Ile	Cys	Leu	Ala	Val	Phe	Pro	Tyr
1				5					10					15	
Gln	Gly	Ser	Ser	Ile	Ile	Leu	Glu	Ser	Gly	Asn	Val	Asn	Asp	Tyr	Glu
				20				25					30		
Val	Val	Tyr	Pro	Gln	Lys	Val	Pro	Ala	Leu	Ser	Lys	Gly	Gly	Val	Gln
				35				40					45		
Asn	Pro	Gln	Pro	Glu	Thr	Lys	Tyr	Glu	Asp	Thr	Met	Gln	Tyr	Glu	Phe
				50			55				60				
Gln	Val	Asn	Gly	Glu	Pro	Val	Val	Leu	His	Leu	Glu	Arg	Asn	Lys	Gly
65					70					75				80	
Leu	Phe	Ser	Glu	Asp	Tyr	Thr	Glu	Thr	His	Tyr	Ala	Ser	Asp	Gly	Arg
				85					90					95	
Glu	Ile	Thr	Thr	Ser	Pro	Leu	Val	Gln	Asp	His	Cys	Tyr	Tyr	His	Gly

100 105 110

Tyr Ile Gln Asn Glu Ala Asp Ser Ser Ala Val Ile Ser Ala Cys Asp  
115 120 125

Gly Leu Lys Gly His Phe Glu Leu Gln Gly Glu Thr Tyr Phe Ile Glu  
130 135 140

Pro Leu Lys Ile Ser Asp Ser Glu Ala His Ala Ile Tyr Lys Asp Glu  
145 150 155 160

Asn Val Glu Asn Glu Asp Glu Thr Pro Glu Thr Cys Gly Val Thr Glu  
165 170 175

Thr Thr Trp Glu Ser Asp Glu Ser Ile Glu Lys Thr Ser Gln Leu Asp  
180 185 190

Asp Asp Asp Lys Arg Pro Pro Thr Asn Thr Pro Glu Gln Asp Arg Tyr  
195 200 205

Leu Gln Ala Lys Lys Tyr Leu Glu Phe Tyr Val Val Val Asp Asn Ile  
210 215 220

Met Tyr Arg His Tyr Lys Arg Asp Lys Pro Val Ile Lys Arg Arg Val  
225 230 235 240

Tyr Glu Met Ile Asn Thr Met Asn Met Val Tyr Asn Arg Leu Asn Phe  
245 250 255

His Ile Ala Leu Ile Gly Leu Glu Ile Trp Ser Asn Arg Asn Glu Ile  
260 265 270

Asn Val Gln Ser Asp Val Gln Ala Thr Leu Asp Leu Phe Gly Glu Trp  
275 280 285

Arg Glu Lys Lys Leu Leu Pro Arg Lys Arg Asn Asp Asn Ala Gln Leu  
290 295 300

Leu Thr Gly Ile Asp Phe Lys Gly Thr Pro Val Gly Leu Ala Tyr Ile  
305 310 315 320

Gly Ser Ile Cys Asn Pro Lys Ser Ser Val Ala Val Val Gln Asp Tyr  
325 330 335

Ser Ser Arg Thr Ser Met Val Ala Ile Thr Met Ala His Glu Met Gly  
340 345 350



595

600

605

Cys Ser Asn Arg Gln Cys Val Asp Val Lys Thr Ala Tyr  
610 615 620

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

ACNCCNGARC ARGAY

15

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

RTAYTTYCKR TACAT

15

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

CAGGACAGGT ACTTGCAGGC CAAA

24

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

ATCGAGTTTT ACGTGGTTGT GGAC

24